DOCUMENT RESUME

ED 373 554 FL 022 354

AUTHOR Davies, Sheena; Parkinson, Brian
TITLE Peer Observation and Post-Lesson Discussion.
PUB DATE 94
NOTE 22p.; For serial publication in which this paper appears, see FL 022 351.
PUB TYPE Reports - Research/Technical (143) -- Journal Articles (080)
JOURNAL CIT Edinburgh Working Papers in Applied Linguistics; n5 p30-49 1994

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Applied Linguistics; *Coding; *Discourse Analysis; Higher Education; Language Patterns; Language Styles; *Language Teachers; Peer Relationship; Speech Acts; *Teacher Attitudes; *Teacher Evaluation; *Teaching Methods

ABSTRACT

This paper describes a collaborative project that sought to combine teacher development and illuminative research. Eight teachers on a general English course, working in pairs, observed each other's lessons and then held discussions, which were recorded and analyzed both quantitatively and qualitatively. Published coding systems were used by the teachers during the observation of their colleague's lesson, while a two-dimensional coding system (covering topics and speech acts) was developed for the quantitative analysis of the post-lesson discussions. Teachers reported that the opportunity to observe their colleagues was very valuable, and that the published coding systems were useful, though at times constraining. The coding system designed to analyze the discussions, though found to be unreliable on the speech act dimension, proved helpful in illuminating patterns of interaction. (Contains 20 references.) (MDM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 373 554

Peer Observation and Post-Lesson Discussion

Sheena Davis and Brian Parkinson (DAL)

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Brian Parkinson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

FL 022 354

PEER OBSERVATION AND POST-LESSON DISCUSSION

Sheena Davies and Brian Parkinson (IALS)

Abstract

This paper describes a collaborative project which sought to combine teacher development and illuminative research. Eight teachers, on a General English course, working in pairs, observed each other's lessons (one lesson per teacher) and then held discussions, which were recorded and analysed both qualitatively and quantitatively. Coding systems were used at two stages: by the teachers during the observation (published systems), and by the researchers, who devised a two-dimensional system (topics and 'speech acts') for the quantitative part of the analysis of discussions. Teachers reported that the opportunity to observe was very valuable, and that the published coding systems were useful though at times constraining. Our discussion analysis system, although unreliable on the 'speech-act' dimension, proved helpful in illuminating patterns of interaction.

1. Background

1.1 Types of observation and ways of doing observation

The observation of FL lessons has a long history, but only in the last 10 - 20 years has it "come of age", with a proliferation of observation systems, articles and textbooks.

We can distinguish four main types of observation, according to observer identity and purpose -

(i) by "experts", as part of teacher "training" and teacher evaluation. This type occurs frequently in pre-service programmes, and is also done by inspectors, head teachers etc.

(ii) by researchers, as part of an attempt to describe and understand classroom events. This may be as part of "pure" research (e.g. Mitchell et al. 1981) or curriculum evaluation (e.g. Parkinson et al. 1981).

(iii) by trainees, observing experienced teachers or each other, presumably in order to learn, crudely speaking, how and how not to teach.

(iv) by practising teachers, observing each other or peers - see Section 1.2.

We can also distinguish three main ways of doing observation, although these distinctions are not clear-cut:-

(a) Unstructured: the observer simply makes a note on what he/she considers salient – sometimes no more than a global impression of the lesson. This type has often been associated with (i) above, and is also very common in (iii). It may indicate lack of thought about observation, or reliance on unarticulated professional experience. The rare cases of genuine ethnographic observation can also be included here, though very different in other ways.

(b) Semi-structured/high inference: the observer has a list of questions to answer about the lesson, and sometimes also boxes to fill in, but these tend to require global decisions and exercise of considerable judgement. This way of doing observation is sometimes criticized for unclear criteria and low reliability, but may be justified when observers (and users of the information) have similar expertise and values.

This way of doing observation has been used for all of types (i) to (iv) above. As an example, we give an extract from a checklist for Practical Tests formerly used in the Diploma for Overseas Teachers of English run by the Royal Society of Arts, reproduced in Mallamah Thomas (1987):

PERSONAL QUALITIES	GRADE	COMMENTS
Personality: Presence, general style		
Ability to establish rapport		
Voice: Audibility, ability to project		

PREPARATION	GRADE	COMMENTS
Lesson plan: balance and variety		
Clarity, limitation and specification of aim		
Suitability of materials and methods for level and type of class		

(c) Structured/low inference: the observer's record of the lesson is compiled according to a detailed rubric or 'coding system' with a pre-set list of categories and guidelines on how often to make a coding. Usually but not always, the category requires less inference and judgement than in (b) above. Examples of coding systems are given later in the paper.

This way of doing observation has been common in type (iii), especially the FIAC (Flanders 1960) and Flint (Moskowitz 1971) systems, and even more so for type (ii) (examples later – see Chaudron 1988 for a review.)

1.2 Peer Observation

From our experience, informal contacts and reading it seems that in most or all kinds of teaching (countries, subjects, class types etc.), peer observation is much less common than observation by someone in 'authority' [(ii) above] or in training [(iii) above]. Many writers on teaching mention peer observation in passing as a sound idea, but in practice time is short and other priorities, and perhaps embarrassment, interfere. Especially rare is peer observation which is systematic in any sense – either

in that it uses a coding system, or even in that there are definite schedules and objectives.

An important exception is the work of Fanselow, who has stressed the value of peer observation for helping teachers in 'generating and exploring alternatives', and has advocated his own FOCUS system for this (Fanselow 1977). In our experience teachers exposed to Fanselow's work are impressed by his general perspective, non-dogmatic and exploratory, but the FOCUS system itself is found complicated and daunting, and other systems, originally designed for research purposes, are preferred. It seems that the FOCUS system has had fairly extensive use in the author's own circles, and perhaps elsewhere in the USA, but the extent of its wider use and usefulness is unclear.

There have, however, been several recent articles stressing the value of peer observation, e.g. Richards and Lockhart (1990), Lockhart (1991), and teacher development through peer discussion e.g. Edge (1992) and Underhill (1992). These are useful in that they provide different perspectives on the topic but only Richards and Lockhart discuss empirical work and there is no detailed description of their findings. Similarly, though the first two articles both mention lesson coding systems - Lockhart says that after deciding what to look for teachers "can either create or choose an instrument which best codes this behaviour" - details of systems and their operation are absent.

The article by Edge describes a framework for peer discussion, this framework, containing nine categories (adapted from Egan) - Attending, Reflecting, Focusing, Thematising, Challenging, Disclosing, Goal-setting, Trailing, Planning - describes and defines the style of interaction and is part of a particular approach which Edge calls "Co-operative Development". Though it was designed for a differently constrained peer discussion, we found it an original and useful, if indirect, input to our work.

2. The Context Of Our Research

The research took place in 1992 on General English classes at the Institute for Applied Language Studies. The General English course is a full time course of 20 hours per week and the students are, in the main, young adults of various nationalities who come for a full term of 11 weeks. The majority study English to improve employment opportunities, a minority to prepare for post-graduate qualifications. Classes at different times of the day focus on the development of different skills.

The teachers are all well-qualified professionals, and all had had previous experience of being observed and observing others. Three had briefly encountered coding systems some years earlier on a master's course in applied linguistics, beyond this, however, none had ever used particular coding systems and no structured system of peer observation has, to date, been set up at IALS.

3. The Experiment

3.1 Objectives

These were stated as follows in the research proposal:

1. To provide teachers with an opportunity to learn by observing other teachers' lessons, receiving comments on their own and discussing issues arising, supported if required by classroom observation literature (e.g. Allwright & Bailey 1991).

2. To inform the academic community on the outcome of this process (research paper) with particular attention to:

(i) the terms in which teachers conceptualise their own and others' lessons
(ii) what is perceived as different
(iii) what is perceived as surprising
(iv) what attracts positive, negative and neutral comment
(v) what use, if any, is made of classroom observation instruments or other help provided
(vi) in what ways, if at all, teachers would like to continue the peer education and self education process

3.2 General Procedure

Eight teachers – all those working on the GE course at the time – were invited to participate in the project, on a voluntary basis, and all agreed. They were put into four pairs, and each member of the pair observed one lesson by the other member for at least one hour (lessons last 90 or 100 minutes). The teachers then had a post-lesson discussion in two parts, one for each lesson, each part to last approximately 30 minutes. The researchers were not present at either lessons or discussion, except in one case where a researcher (SD) was one of the teachers. The lessons were not recorded but the discussions were, and the research was conducted on the understanding that discussions, not lessons, were the main focus of investigation. It was also stressed that the research was non-evaluative in a double sense: observers should not 'judge' the lessons, and researchers would not judge the discussion comments: the purpose was mainly one of professional development exploring whatever issues were of interest.

Tapes of the discussions were transcribed (by research assistants who were experienced EFL teachers), and these transcripts (checked with the original where necessary) constitute the main part of our data. They are supplemented by two minor data sources:

(ii) the completed coding sheets (see Section 4.1 below) used by the observers and
(iii) a post-discussion questionnaire (see Section 4.2 below).

The teacher/observers were given a selection of recognised coding systems, with background information (see Section 3.3), and asked to select one of these before observing and use it during observation. In addition, they were asked to make notes

6 BEST COPY AVAILABLE

of anything observed which was not covered by the system but which seemed surprising, interesting, etc. (cf. research questions, of which observers had a copy). To facilitate this, observers were provided with a three-page coding sheet with space for both system-based comments and open-ended comments.

The observers were not expected to use the systems 'properly', i.e. to make exhaustive coding using exact definitions. This would have been impossible without extensive training. Instead, they were asked to use them as a basis for entries which indicated the main patterns of the lesson. It was stressed, however, that something more than a general impression was required, and that sequential, timed coding should be attempted.

In the post-lesson discussion the teacher observers were asked to discuss the lessons in whatever way they felt useful. This could be, but did not have to be, based partly on the coding sheets.

3.3 Coding Systems Offered and Used

The following is a brief summary of the systems offered for observer use:

(i) 'BIAS' system (Brown 1975). Indicates whether teachers or learners are speaking, and whether they are lecturing, questioning, responding etc.

(ii) Bowers' system (Bowers 1980). Looks at social functions of classroom language – presenting, directing, organising, eliciting, evaluating, responding, socialising.

(iii) 'COLT' system (Frohlich et al 1985). A high-inference system judging presence of communicative features in interaction.

(iv) Allwright's error treatment system (Allwright 1988). How teacher treats errors.

(v) Chaudron's system (Chaudron 1977). As (iv) but more complex.

(vi) Embryonic system (Long et al. 1976). A list of 17 'pedagogical moves' e.g. initiates discussion, summarizes, provides example; 13 'social skills', e.g. interrupts, contradicts, encourages, jokes; 14 'rhetorical acts', e.g. predicts, hypothesizes, deduces, negates.

(vii) Pica and Doughty system (1985). For the analysis of group work, and how students check mutual understanding.

(viii) Allwright's turn-taking system (1980). Taking turns, 'stealing' turns, offering the floor to other learners and so on.

Only the first three of these systems were actually used.

3.4 Data Analysis

This focussed almost entirely on the discussion (other minor sources – coding sheets and questionnaires – are briefly discussed in the results section). Analysis was mainly from the transcripts, but the original tapes were listened to where necessary.

To analyse the transcripts, we devised a two-dimensional coding system, covering (i) the types of topics discussed and (ii) the ways in which teachers interacted, and what we perceived as the underlying speech acts. Our categories were largely 'post-hoc', i.e. created to cover what we found in the transcripts, but we attempted to keep in

mind our research questions and make only those distinctions necessary to answer these.

The system was devised jointly by both researchers, and several joint codings were attempted in order first to improve and then to measure inter-coder reliability. Due to the small amount of data, a rigorous reliability study was not possible but, after coding, a final inter-coder reliability check was made and the results are reported in Section 4.3.1 below.

The categories used, with examples, were as follows:-

A. Topics

Number	Topic	Example
1	Facts about students e.g. nationality, level, age, history	"He's been here for ages and he's going to be here for even further ages"
2	Observation system used by coder	"That's the, erm the one that is supposed to measure the social functions of the language that goes on"
3	Expectations	"Had you got any preconceived ideas about what would be happening?"
4	Relation between expectations and events	"No well I suppose I imagined probably, about what I saw. I mean some input and some practice and some real communication"
5	The observed lesson (general)	"So that's why they had this sort of check list they were using"
5.1	The observed lesson (learner behaviour)	"He was making a face [ ] he wasn't seeking confirmation"
5.2	The observed lesson (teacher behaviour)	"And at one point you sort of broke off everything and said 'You look puzzled' "
6	Other lessons by same teacher (general)	"We had been doing quite a lot of work beforehand on discussion techniques"
6.1	Other lessons by same teacher (learner behaviour)	"He doesn't often do that actually"
6.2	Other lessons by same teacher (teacher behaviour)	"In that kind of situation I don't, er, if they ask me a question I just turn away"
7	Other lessons by observer (general)	Similar to 5 and 6
7.1	Other lessons by observer (learner behaviour)	Similar to 5.1 and 6.2
7.2	Other lessons by observer (teacher behaviour)	Similar to 5.2 and 6.2
8	Other lessons or lessons in general (general)	Similar to 5 and 6
8.1	Other lessons or lessons in general (learner behaviour)	Similar to 5.1 and 6.1
8.2	Other lessons or lessons in general (teacher behaviour)	Similar to 5.2 and 6.2
9	Linguistic theories concepts	"It's amazing how much, well er, I won't say students, I mean anyone can, how much you can read without actually taking any of it in."
10	Materials and syllabus	"Was that something that came up in the textbook?"
11.1	Personal feelings (general)	"I was genuinely surprised"
11.2	Feelings about being observed	"... when you've sort of er being watched by your peers as it were, you do feel a certain, that you are being judged [ ] it is a bit nerve racking..."
11.3	Feelings about discussing lesson	"I don't want to sort of make any evaluative comments or judgements on the thing"
12	The English language	"But I mean what is the actual dictionary definition of 'authoritative'?"

8 BEST COPY AVAILABLE

13.1	The project in general	"Well it seems to me that to get the most value from anything like this, viewing each other once is just nothing like enough."
13.2	Observation in general	"I wonder how far what one learns from observing is actually useful when you're actually teaching. Maybe it is in the planning stage."
14	Other	What did you say?

B. Interaction patterns underlying "speech acts"

Label	"Speech Acts"	Examples
BC	Back channelling	"Ah, I see. Right, right."
SY	Sympathising	"This is a big problem isn't it, that everybody has."
SP	Speculating	"Maybe she only talked a lot the second time because she was [playing another role]."
QN	Question (neutral)	"Was it difficult to try to use it?"
QC	Question (challenge)	"But why did me being there stop you doing it spontaneou[sly]?"
INF	Informing	"After you left they started discussing again."
OP	Opinionating	"I think X is very much a sort of actor anyway."
EV +	Evaluation (positive)	"It was good that it was slipping into real discussion."
EV -	Evaluation (negative)	(i) "[My] board work was pretty awful." (ii) "The only thing was, the only thing was, that it was, they, the particular constructions with the particular prompts did seem to be causing some rather odd constructions."
IE	Inviting evaluation	"Was there anything else that surprised you, that you wanted to ask me why you did something? 'Cos it's very useful for me as well to get sort of, a reaction of some kind."
A	Adverting or semi-phatic	"You reminded them about erm what they had done the previous week."
J	Justifying	"I mean the point of that exercise was to see whether they had actually got their heads round the, erm, the distinction."
SG	Suggesting	"If they'd had a piece of paper with it written down it probably would have been easier."
AG	Agreeing	"Yes, Anna said that a couple of times, yes."
Other	Other	"Oh, maybe."

N.B. These "speech acts" must be interpreted in their interaction with different topic categories, and with who (teacher or observer) is speaking. For example, commonsense suggests that evaluating one's own lesson, one's interlocutor's lesson and a third party's lesson are three very different kinds of speech acts. The grand total for each "speech act" may thus sometimes be less illuminating than totals for individual combinations of topic and "speech act".

4. Results

4.1 Information from Lesson Coding Sheets

Six of the eight teachers handed in their coding sheets. (They were not obliged to do so since the focus was on the post-lesson discussion.)

When analysing the data, we asked ourselves the following questions:

- how long did they keep on coding?
- did they use their own categories or those of the 'system'?
- what was the balance between the 'system' observations and the open-ended observations?
- what type of comment/information was in the open-ended column?

- did they use the third checksheet (for general comments)?

All the teachers coded the lesson events for the full hour, following the time segments. The use of the categories of their chosen systems varied from full use of all the categories (by two teachers using the BIAS system and one using BOWERS) to partial use of the categories (by two using COLT and one BIAS). Most teachers gave some illustrations or explanations of these categories in the system column. For example: *"Sociating - rearranging bodies"* (BOWERS); *"S [= Silence] because focussing their minds & individually formulating questions"* (BIAS).

In the open-ended column, most comments were descriptions of lesson events - a combination of more illustrations of the system categories and descriptions of other events focussing on teacher or student behaviour, or examples of language used, such as: *"Qs on OHT on topics of film"; "E. demonstrate swagger by walking"*. One sheet also included several positive evaluative comments: *"most Ss participating well"; "T good at involving any S who is reticent left out"*.

The balance of observations between the two columns seemed fairly even, and in most cases there was nothing substantially different between the two, apart from one coding sheet in which the open-ended column consisted mainly of diagrams of the teacher-student interaction. This reflected the personal interest of the teacher concerned, an interest she elaborated on in both the post-lesson discussion and the general checksheet 3.

Checksheet 3 gave observers the opportunity to make any further comments and was submitted by only four teachers. They each used Sheet 3 for further descriptions of lesson events, but two also made comments and posed questions about the system and what they were trying to do. For example, *"Is communication genuine if T clarifies language for S?"; "Difficult to concentrate on different levels of communication e.g. interpersonal, formal, plenary"; "Prediction of a lesson format v. difficult if not impossible"*. These issues were later raised in the post-lesson discussion between the teacher and the observer.

4.2 Information from Questionnaires

Tutors were asked to fill in a questionnaire to supplement the data from the discussion. The questions were open-ended (see below) and anonymity was guaranteed.

1. Why did you choose the observation system(s) you did? Was it/were they a help or a hindrance?
2. What, if anything, did you find useful or valuable about
   (a) observing another's class?
   (b) being observed?
   (c) discussing the lessons?
3. Did you feel any constraints in the post-observation discussion?
4. How, if at all, would you like to follow this up?

The responses, though varied, show some measure of agreement. Not surprisingly perhaps, four out of the eight tutors stated that they had chosen the system because it

reflected an area of interest, while the others selected a system for its apparent manageability for real-time coding and its "degree of teacher-friendliness".
"I chose a system that seemed best suited to look at teacher-student interaction" (BOWERS); *"It seemed reasonable to handle within the constraints of timing and is an area of personal interest."* (BIAS)

With regard to its usefulness, there were only two unqualified answers - one found it useless, the other a positive help by providing a systematic framework. The remaining tutors found it of limited usefulness - a help initially by giving a guideline or focus but then becoming a hindrance because it was too restricting or difficult to use, or inappropriate for the actual classroom events. *"A good guideline, but perhaps too narrow."*; *"On the whole it was a help but sometimes it was difficult to observe other useful and interesting events because of the focus of the observation system."*

All the tutors, however, were positive about observing another's class and there was strong agreement in the reasons given, such as: *"interesting and informative to watch how someone else deals with a topic ..."*; *"makes you review your own teaching methods"*; *"it is always useful to pick up new ideas from other teachers"*

The comments on being observed echoed the comments above, for example: *"having another person's opinion about a problem"*; *"makes you think a little more about what you do and say, and how much you say"* and also included the opinion that *"As I believe in team teaching as a useful method of teaching certain types of classes and students, I think it is helpful to be observed and to be able to feel comfortable and not in any way inhibited by the presence of another teacher."*

Similarly, the post-lesson discussion attracted positive comment - again, mainly echoing the points made above but two teachers made similar observations about interpretation and perceptions: *"interesting to see whether what the teacher felt had been important/salient corresponded to what the observer felt"*; *"I found it interesting, and quite surprising sometimes, how an observer interpreted what I did"*

Only one person found the experience of being observed and the post-observation discussion nerve-wracking; nor did this teacher like the discussion immediately after the lesson, preferring (in retrospect) to have had more time to reflect and be more analytical.

There was a high level of unanimity among the responses to the questions on feelings of constraint in the post-observation discussion. Five tutors commented that they would not want to say, or tried to avoid saying, anything critical about their colleague's lesson because it was a peer situation. This does not mean that they felt there had been something negative to say, only that they would not have wanted to say it if there had: *"If I really didn't like something I saw or had a negative criticism (which was not the case in the one observation) I feel that this set-up would not encourage me to say anything."*; *"I wouldn't want to say anything negative (even if I had something to say) as it was a peer situation and as such should be supportive and non-judgemental."*

This conforms with the emphasis on the non-evaluative nature of the research project in general and the lesson observation in particular (see 3.2 above) but it is interesting that such a feeling should be perceived as a constraint. Only one person implied that

there were negative points made about the observed lesson. Other constraints mentioned were the artificiality of talking into the microphone (initially), and talking in a more "formal" way rather than a chat over a cup of coffee.

With regard to possible follow-up, the points made tended to fall into two main categories:

(a) introduce a system of frequent observation as routine teacher development, not as a research project (five teachers mentioned this)

(b) if the project is repeated, to have a more focussed system on topics agreed on by participating teachers (two teachers).

4.3 Introduction to Discussion Coding Tables

The following tables give the overall coding totals for observers and for teachers. They show the number of occurrences in each category, and each combination of "topic" and speech act category.

Table 1
Observer Moves - Grand Total

	BC	SY	SP	QN	QC	INF	OP	EV+	EV-	IE	A	CN	I	SG	AG	Other	TOTAL
1 Facts about students	6	2	1	15	4	1	3				3				2		37
2 Coding system	2			1		38	4	1	11				1			1	59
3 Expectations				1		4			1								6
4 Relation expect event				1	1	5			1		1						9
5 Observed lesson (general)	22	5	3	9	2	21	5	28	6	1	16	2	1	3	1	2	127
5.1 Observed lesson (SS)	7	1	5	7	4	28	26	23	7		37		5	2	12		164
5.2 Observed lesson (T)	1	2		8	5	17	9	37	10		31		9	6	1		136
6 T's lesson (general)	10			9	4	2	2										27
6.1 T's lesson (SS)	7		1	5			1										14
6.2 T's lesson (T)	3	1		8			3						1				16
7 O's lesson (general)		1		2		12								1			16
7.1 O's lesson (SS)						1											1
7.2 O's lesson (T)						4											4
8 Other lessons (gen.)	2	7	1	3		6	5				1			1	3		29
8.1 Other lessons (SS)	2		1			3	8				1				3		18
8.2 Other lessons (T)	3	1	2			3	3							1			13
9 Theory			2			4	6										17
10 Material	6		1	9	2		1	1			3						23
11.1 Feelings (general)	1			1		10											12
11.2 Feelings (observed)			4	2		8	7					1			2		21
11.3 Feelings (discussing)						5			1								6

12

BEST COPY AVAILABLE

12	English	2		1	4	1	3	1				1	1			1		15
13	Project/obs (general)			1			1	4										6
14	Other				2		1									1	2	6
	TOTAL	74	20	23	87	23	174	88	90	37	1	94	4	17	14	26	5	777

Table 2
Teacher Moves - Grand Total

		BC	SY	SP	QN	QC	INF	OP	EV+	EV-	IE	A	CN	J	SG	AG	Other	TOTAL
1	Facts about students	2					35	4	2			1				4		48
2	Coding system	6		1	6		1	3		1		2	2	1		1		24
3	Expectations				2		1											3
4	Relation expect/event				2		9		1		1							13
5	Observed lesson (general)	16			6		28	9	8	4	8	16	3	31		8	1	138
5.1	Observed lesson (SS)	3	1	4	5		29	20	13	5	1	17		6		11	1	116
5.2	Observed lesson (T)	10		3	2		7	1	6	4	9	4	9	33	3		1	92
6	T's lesson (general)	1			1		35	5		1	1		1	3		2		50
6.1	T's lesson (SS)			1	1		42	11	3	3		1		1				63
6.2	T's lesson (T)						35		1		1		1	14		1		53
7	O's lesson (general)						2					1						3
7.1	O's lesson (SS)																	0
7.2	O's lesson (T)																	0
8	Other lessons (gen.)	2		1			2	8			1	2		1		3		20
8.1	Other lessons (SS)	1	1				3	9	1	2		1				3		21
8.2	Other lessons (T)			2			3	7					1		1			14
9	Theory	1					2	5								1		9
10	Material	1		1			18	1	1	2		2	2	1				29
11.1	Feelings (general)			1	1	1	6	1					1					11
11.2	Feelings (observed)	2		4		4	4	4								2		20
11.3	Feelings (discussing)																	0
12	English	1	1	1	1		2	3								1		10
13	Project/obs (general)						1	2										3
14	Other	1			3		4	1								1		10
	TOTAL	47	3	19	30	5	269	94	36	22	22	47	20	91	4	38	3	750

4.3.1. Inter-coder reliability

As an inter-coder reliability check, half of one discussion, chosen at random, was coded by both researchers. We chose not to give a single percentage figure or set of figures for reliability, considering this to be meaningless as it could be calculated in many different ways. The statistics seem to show, however, that our 'topic' figures have some inter-subjective status: agreement is as high as could be expected in a small, exploratory study. The "speech act" agreement, on the other hand, is disappointingly low, and means that our figures on this dimension must be treated sceptically, though they can give some general indication of how teachers behaved.

As far as we can tell, our difficulties were not mainly to do with conceptualising the categories: we seemed to agree on what we meant by informing, evaluating etc. - but on interpreting specific utterances: illocutionary force is often not totally explicit, and it is well known that it may be difficult for third parties, and even for participants, to recognise it with certainty.

4.4 Annotated selection of teacher comments

This section aims to supplement the quantitative information on the discussion with a selection of teacher comments, organised by topic, which seem particularly relevant to the research questions. We have omitted hesitation devices and repetitions as far as possible, except where they seemed important to an understanding of the interaction, but we have not achieved total consistency in this, nor in our indication of minor abridgement. Comments by both observer [O] and the teacher [T] being observed are included.

4.4.1 Comments on use and usefulness of coding system

As already indicated, we asked teachers to be guided by the system rather than to attempt exhaustive coding of every utterance, and most did exactly that. Nonetheless, several offered reasons (in the discussions) for their limited use of their chosen system, often coupled with comments on its adequacy or otherwise:

"I was trying to use ... Brown's Interaction Analysis (BIAS) ... When I looked through the ones on offer, it seemed to be quite interesting. It was in the area that I wanted to consider. But ... it's very difficult to actually stick to one of these systems exactly, because there aren't enough categories. Or these categories given here are not exactly relevant to the type of lesson you were doing. I would have liked to include more categories because some of these ... sound negative e.g. 'teacher lectures'. To me that has a negative overtone that the teacher's jawing away all the time. Then it says "explains" and "directs" - well yes, you did that a few times, on the sheet I've put down "T." in various situations, but it was very short, succinct and straight to the point ..." [O]

We would not presume to reduce such comments to a one sentence summary, but the main points are perhaps (i) the "ethnographic" one that pre set systems fail to capture the richness of experience, (ii) mismatch between system and modern methods and (iii) distrust of evaluative labels.

4.4.2 Comments on effect of observation

As usual in studies of this type, the question arose of the effect, if any, of observer presence on teacher behaviour and learner behaviour. Effects on the teacher were only rarely mentioned. *"I'm always a bit nervous though, I think you tend to be, no matter what you say you still think you're being watched and evaluated."* A possible effect on learners was mentioned slightly more often. *"They're usually quite a lively class specially first thing in the morning ... so it might have been the fact that you were here that might have had an influence."* The majority view, however, seemed to be that the effect of observation was small, e.g. *"No, I don't think they were affected. X has been in to watch them for a lesson, and people pop in and out, and I film them ... so I don't think they're bothered too much by other people"*

4.4.3 What attracts positive comments

Although all observers showed a positive and supportive attitude towards the teacher and the observed lesson, they did not find it necessary to make frequent comments of direct praise. Interestingly, many evaluative comments related to students rather than the teacher, although some of these might be interpreted as indirect praise of the teacher. *"They seemed to be working very naturally together"*, *"They really did seem to get on well together, that was nice. I mean there was quite a lot of er, I was quite surprised, there was quite a lot of touching actually between them ... you know, playful slaps on the and that kind of thing"*, *"I mean there was a tremendous enthusiasm there wasn't there, they were really sort of getting in there"*

An example of direct positive comment on the teacher was *"I felt that the lesson was very carefully crafted."* Less clear as a speech act, but also classified as positive comment on the teacher, was the comment *"it was quite interesting the way you were anticipating that they would have problems, you'd be much more careful about how you communicated"*

An area in which reassurance was sought in several discussions was that of learner involvement. Teachers said that they, often or on a particular occasion, were unsure if activities went on too long or not long enough, if some students were bored etc., as they were unable to monitor all learners all the time. Observers replied positively in all cases - as far as they could see pacing and involvement were satisfactory, though in some cases it was agreed that differences in learner level etc. meant there was no ideal solution.

4.4.4 What attracts negative comments

It is a general rule of human interaction that negative comments are made far more often about (especially absent) third parties, far less often about the speaker, least often about the addressee, so it is no surprise that most comments in the 'negative evaluation' category referred to students, e.g.: *"A tends to wander off into his own little world quite a lot"*, *"B can be quite aggressive"*, *"There's one or two who don't like being corrected ... seem to lose face a bit"*

It is striking, however, that such comments usually refer to one student, most of the rest to small sub-groups, very few to whole classes. The teachers generally seem to

BEST COPY AVAILABLE

have positive attitudes to their classes, and even the criticisms of individuals may usually be interpreted as relating to the difficulty of doing certain kinds of work with them rather than as outright hostility.

Self criticism was only a minor feature of the discussions: although not infrequent it was typically brief, and in the nature of an aside: *"My board work was pretty awful."*

Criticism of the teacher is infrequent and almost always mitigated or qualified. We do not suggest, however, that observers were withholding or excessively 'toning down' negative comments for reasons of tact. One can never know, but the evidence (see 4.2) and our impressions suggest that the teachers did feel generally positive about the lessons and about each other. But the difficulty of expressing even minor and constructive criticism needs to be borne in mind.

4.4.5 What is perceived as surprising?

Observers seemed to find the lessons interesting, but rarely surprising. This was perhaps to be expected as there is a culture of cooperation within General English and a lot of informal discussion about what happens in classes. For an exceptional 'surprise', see 4.4.6.5.

4.4.6 Specific issues of methodology and learner behaviour

Under this heading, we consider methodological and similar issues which arose, some in only one discussion, some in most or all discussions.

4.4.6.1 Genuine communication versus practice of forms

The first of these, the proper role for 'communication' activities and practice of forms, and indeed how to distinguish between these, was addressed in some form by most of the participants, most extensively when the COLT system had been used.

"... what we're trying to achieve in the afternoon classes, we're trying to promote real communication but we have to give practice activities ... specially advanced level ... they're not like practising one particular function, it's usually some kind of a strategy ... so the fact that you give them a task in which you want them to practise this but the task itself is not too constraining, is that genuine communication or not?"
(and later)
"When you were eliciting things from them, and although you didn't know what their responses would be, is that genuine communication or not?" [O]

Learners in some classes had been given a range of exponents for expressing opinion, taking the floor and related functions, and when they practised these there was sometimes doubt about whether they had to express genuine opinions. *"It was very interesting to try to gauge at different points to what extent they were having a genuine discussion [or] just practising language very consciously and it seemed to slide very much backwards and forwards between the two."* [O]

4.4.6.2 Error treatment

This topic arose with similar frequency to the last, and at greater length. Space prohibits a full account of conclusions reached, but we quote from two discussions to give a flavour of the ideas expressed. "... *it was quite interesting to know through the progression of the lesson whether it's certain activities that lend themselves to a certain kind of error correction, in which case it's quite important to plan one's lesson.*" [O]

In the following example, an observer reflects how different kinds of class at different times of day (see Section 2) require different attitudes to error. "*I expect it to be more teacher-directed than student-centred which in fact happened and I also expected there to be more error correction, you know, than in a fluency class or something like that. ... Error corrections and pronunciation checks and things like that ... so, my preconceptions were justified so it wasn't surprising from that point of view, you know, given the kind of lesson it was.*"

4.4.6.3 Teacher as model (articulation)

This is an interesting example of a one-off comment which derives from an important issue on the observer's personal agenda. No criticism seems to be implied, only wondering aloud. "*... it was quite interesting - and this is something I always worry about, you were articulating very carefully when you were speaking to the students. I do that. I just wonder are some teachers doing it, you know as a sort of, saying it in a very fluent native speaker way and not articulating.*" [O]

4.4.6.4 Dominance in group work

The observer in one discussion chose to investigate whether particular learners were dominant in group work whilst others said little: this is of course a common area of both system-based and open-ended observation, promoted especially by Allwright. The conclusion was that in this case no-one was obviously dominant, and no-one excluded, but an interesting pattern was noted. "*It was A and B who were doing most of the talking, and then the blonde woman, well her role seemed quite interesting. ... she seemed to be slightly also taking on the teacher's role in that she sometimes, I think she once or twice corrected people or provided words that somebody else was looking for. ... they seemed to ... look to her ... as a sort of linguistic consultant.*"

4.4.6.5 Learner independence

One discussion was different from the others in that, in the observed lesson, the teacher had adopted a range of procedures which both teacher and observer perceived as somewhat unusual or individualistic, and which had an explicit rationale of encouraging learner independence. Thus, for example, the class had watched TV programmes, chosen by themselves, aided by their own lists of predictions/questions rather than teacher-devised worksheets, in order to show them a way to benefit from watching normal TV at home; and many questions about vocabulary, even requests for guidance on 'the best way to learn', were turned back for the students to answer themselves. In the discussion, the observer generally praised this approach ("*I love the way you did that*"), but also voiced doubts, e.g. "*I'm just wondering what happens*

if nobody knows". The teacher elaborated the thinking behind the approach, but modestly disclaimed any definitive answers and accepted that the approach would not always work, even giving an example of a 'failure' after the observer had left the lesson.

As mentioned above (4.4.5), not much in the observed lessons appeared to be perceived as surprising, and this discussion seems to be the only case within the discussions of something being described as radically new.

Other teachers too were very conscious of the learner independence issue, and provided examples in their actions as teachers or their suggestions as observers. One teacher, for instance, had invited a student to explain something to another student, and justified this as follows: *"Well I like to do that because, rather than me just saying 'It means this', which in that state I'm becoming the dictionary and I'm trying to wean them off these wretched dictionaries so it's better if they're trying to ..."*

The next example shows the same desire for learner independence, but a less satisfactory reality: *"what's really unfortunate is even if they are in groups, they'll still kind of try and turn round and involve me and sort of call me in you know."*

Finally, the 'learner independence' area provides one of the few examples in the post-lesson discussions of a detailed suggestion from an observer: *"I noticed that when he'd maybe finished he just sort of tended to, he'd obviously, what was his name, A had obviously finished and you think well, he could have been encouraged more to help, or maybe, I don't know, maybe he didn't like helping his friend, you know, and so he maybe, getting him to explain 'Why did you put this?', 'Why did you put that?'"*

4.4.7 Feelings about doing observation

The general idea of observing each other's lessons seemed to be viewed very positively. As regards the use of systematic coding systems, or any kind of systematic peer observation, comments were more cautious - still positive, but aware of the difficulties and the need for more experience. Sample comments: *"I suppose it's very significant that when teachers sit down to talk about a class, the mode that you tend to slip into is evaluative, even if it's very encouragingly, because most often any time you're actually paid to sit down and talk about someone else's class it's because you're examining them or you're helping them towards an exam, or you're seeing how it should be done from somebody who has more experience."* [T], *"you're switching into observer mode for a particular thing and you're not aware of of all the other things that are going - you're looking for X and there's A to Z going on around you because you're looking for X"* [O], *"I think as practising chalkface teachers we're very interested in how things work - and classroom management techniques and skills and so on ... it's relevant to this observation thing that we're doing that you tend to get caught up with those pedagogical issues"* [O]

5. Discussion And Conclusions

5.1 Discussion of Transcript Codings: 'Topic' Dimension

The following general features may be noted:

BEST COPY AVAILABLE 18

(i) In all discussions for both speakers the observed lesson (category 5, including 5.1 and 5.2) is, not surprisingly, the main topic.

(ii) Within this category the emphasis varies, some discussions emphasising learner behaviour, some teacher behaviour, some more evenly balanced. On the whole, however, learners were discussed more than teachers. This may reflect both the current climate of opinion within EFL and the perceived usefulness of another perspective on learner interaction.

(iii) The second most consistently frequent category - more than 10 codings in each of six discussions - was category 6 (including 6.1 and 6.2), i.e. what the observed teacher did in other lessons, including parts of the observed lesson before and after observation. The felt need to put observed events in context is unsurprising.

(iv) The only other categories which even approached this frequency - each having more than 10 codings in each of four discussions - were category 8 (including 8.1 and 8.2), i.e. lessons in general, and category 2, the coding system.

(v) Only three other categories ever exceeded the '10 codings' threshold, and these only in one discussion each, reflecting a particular focus of that discussion. These were: The teaching material (category 10); Feelings about being observed (category 11.2); The English language (category 12).

(vi) There was a tendency for similar topics to be discussed in each of a pair of discussions with the same participants.

At this point it may be worth repeating that the ethos of the project was non-evaluative, and we shall not presume to suggest that high or low use of any topic is 'good' or 'bad'. The teachers discussed what they identified as worth discussing, exactly as intended. If one were planning a more extensive and structured peer observation study, however, one might wish to look for ways, over a long series of discussions, of ensuring a wide and systematic topic coverage including both teacher and learner behaviour.

5.2 Discussion of Transcript Codings - 'Speech Act' Dimension

The first point to note here is a blurring of the role distinction which might have been expected: for example, not only the observer but also the teacher sometimes evaluates what happened in a lesson, and not only the teacher but also the observer justifies it. This perhaps indicates a high level of mutual supportiveness with the group, with strong desire to convey feelings of solidarity and emphasize the shared features of experience.

Beyond this, we were struck by the wide variation of speech act frequency in the discussions: the differences between the discussions are far more salient than the similarities. For example, two discussions, with their high frequency on the part of the observer, of questioning and 'adverting or semi-phatic' (i.e. mentioning something known to both speakers), seem to be a very different kind of speech event from other discussions, where these categories are much rarer. (Other untypical figures, such as the high 'back-channelling' rate of one observer, may reflect individual speech styles.)

What is common to all discussions, however, is a fairly high rate of 'evaluating' acts by the observer (sometimes, only late in the discussion and after frequent invitations to evaluate by the teacher.) Most of these comments were on the observed lesson, although other topics (materials, students, other lessons) were sometimes evaluated. All teachers made one or more evaluative comments on their own lessons, though sometimes not many and always fewer than the observer. Evaluations were in general positive, and any negative comments were tentative and qualified.

The frequency of evaluative comments seems to confirm that peer observation is always likely to be a partially evaluative process, and that, although one can mitigate this by emphasizing non-judgmental aspects, one cannot, and perhaps should not want to, eliminate it entirely.

5.3 General Conclusions

Our first, and unsurprising, conclusion is that the peer observation was worth doing. Teachers did appreciate the opportunity, surprisingly rare in most professional lives, to observe another teacher's lesson and to be observed without any context of evaluation or bureaucratic requirement.

Second, and less certain a priori, the coding systems seemed to be of some value. Every professional in the area of systematic observation knows that all existing systems are far from ideal, sometimes difficult to apply - especially to classes taught by modern methods - and not always yielding insights, and the observer comments (see especially 4.4.1) amply confirmed this. Nonetheless, observers did persevere with systems, often filling in coding sheets very fully - perhaps more fully than we had expected - and showing great resourcefulness in taking the systems as a starting point for more open-ended comments on topics of major interest.

Our third conclusion relates to the low inter-coder reliability on our 'speech-act' dimension. It seems either that we are unusually incompetent in recognising illocutionary force, or, more probably, that what is said in discussions of this type is even more polysemic (polypragmatic?) than one would generally suppose. This could be a fruitful area for further research.

Fourthly, despite these uncertainties of pragmatic detail, the general goodwill, enthusiasm, mutual supportiveness and professional commitment of the teachers was very much in evidence, as was the structured and principled nature of the curriculum which they were implementing.

Fifthly, it seems that further peer observation, although not an urgent priority, would be of potential benefit to the course, to those of the eight teachers still working on it and to others who have replaced some of them. As usual in such research, any second round of observation could profitably be made slightly more selective and structured, building on the findings of this pilot study. Post-lesson discussion could be similarly guided, with a wide and systematic topic coverage including both teacher and learner behaviour.

Acknowledgement

We would like to thank the teachers who took part in this project - Wendy Ball, Cathy Benson, Sue Fraser, Marie Gilroy, Liam Rodger, Paul Snookes and Richard Warner.

References

Allwright R. 1980. 'Turns, topics and tasks: patterns of participation in language learning and teaching'. In D. Larsen-Freeman (ed.) Discourse Analysis in Second Language Research. Rowley, Mass.: Newbury House. 165-187.

Allwright D. 1988. Observation in the Language Classroom. Harlow: Longman.

Allwright R. and K. Bailey. 1991. Focus on the Language Classroom. Cambridge: CUP.

Bowers R. 1980. 'Verbal Behaviour in the Language Teaching Classroom'. PhD Thesis, University of Reading.

Brown G. 1975. Microteaching. London: Methuen.

Chaudron C. 1977. 'A descriptive model of discourse in the corrective treatment of learners' errors'. Language Learning 27: 29-46.

Chaudron C. 1988. Second Language Classrooms: Research on Teaching and Learning. Cambridge: CUP.

Edge J. 1992. 'Co-operative development'. ELT Journal 46.1: 62-70.

Fanselow J. 1977. 'Beyond Rashomon - conceptualizing and describing the teaching act'. TESOL Quarterly 11: 17-39.

Flanders N. 1960. Interaction Analysis in the Classroom: a Manual for Observers. Ann Arbor: University of Michigan.

Frohlich M., N. Spada, and P. Allen. 1985. 'Differences in the communicative orientation of L2 classrooms'. TESOL Quarterly 19: 27-56.

Lockhart C. 1990. 'Co-operative teacher development: new observations on observation.' Perspectives 2: 43-57.

Long M., L. Adams, M. McLean and F. Castaños. 1976. 'Doing things with words - verbal interaction in lockstep and small group classroom situations'. in J. Fanselow and R. Crymes (eds.) On TESOL '76. Washington: TESOL: 137-153.

Malamah-Thomas A. 1987. Classroom Interaction. Oxford: OUP.

Mitchell R., B. Parkinson, and R. Johnstone. 1981. The Foreign Language Classroom: an Observational Study. Stirling Educational Monographs No. 9. Stirling: Department of Education, University of Stirling.

Moskowitz G. 1971. 'Interaction analysis - a new modern language for supervisors' Foreign Language Annals 5: 211-221.

Parkinson B., R. Mitchell, and D. McIntyre. 1984. An Independent Evaluation of 'Tour de France'. Stirling Educational Monographs No. 11. Stirling: Department of Education, University of Stirling.

Pica T. and T. Doughty. 1985. 'Impact and interaction in the communicative classroom: a comparison of teacher-fronted and group activities'. In S. Gass, and C. Madden, (eds.) Input and Second Language Acquisition. Rowley, Mass.: Newbury House: 115-132.

Richards J. and C. Lockhart. 1991. 'Teacher development through peer observation' Perspectives 3/2: 1-9.

Underhill A. 1992. 'The role of groups in developing teacher self awareness'. ELT Journal 46/1: 71-80.